

Fitting a Line to Data

Learning involves what psychologists call *dissonance*: Unless people see that they don't understand, they can't really learn. Learning an idea involves adjusting your concepts so that the new idea fits. Because we want our students to understand, we might rush to explain concepts rather than let students come to realize on their own that they don't understand. Try to decrease your student's frustration with not understanding a concept right away. Show how your own curiosity allows you to welcome initial questions and confusion as an opportunity to learn.

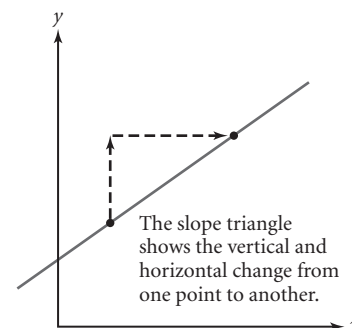
Content Summary

Chapter 4 builds on the ideas of Chapter 3. So students may find that they don't understand those earlier ideas as well as they thought they did. In this chapter, the notion of steepness gets formalized (more strictly defined), and your student learns how to derive another form of linear equations. Challenge your student to be patient as he or she works to understand how to make predictions from data points that appear somewhat linear but don't all lie on any single line.

Slope of a Line

In Chapter 3, students learned that the equation $y = a + bx$ represents a line that goes through point $(0, a)$ and climbs b units for each unit across. So, b measures the steepness of the line. It may be difficult to measure how much a line climbs in one unit; it's often easier to find how much it climbs over a larger number of units, and then divide. The result is the amount that it climbs for each unit. This number b is the line's *slope*, the mathematical term for *steepness*.

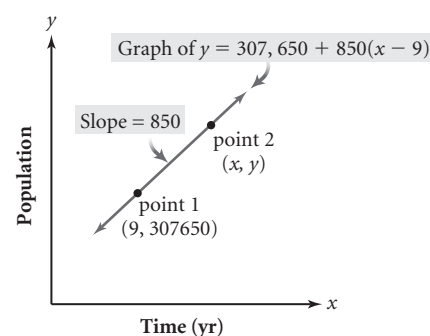
To calculate the slope of a line, students draw a *slope triangle* to find how much the line climbs over some convenient horizontal distance. From this idea comes the formula for the slope of a line through two points; the slope is the change in vertical distance divided by the change in horizontal distance.



Point-Slope Form of a Linear Equation

One reason the text emphasizes the intercept form of linear equations ($y = a + bx$) in Chapter 3 rather than the slope-intercept form ($y = mx + b$) is to convey the idea that linear growth starts at a and climbs b units for each unit change in x . Another reason is that, if the growth starts at values of x other than 0, the equation generalizes naturally. For example, if the growth begins at point (x_1, y_1) , the equation is $y = y_1 + b(x - x_1)$. This is the point-slope form of a linear equation.

In Lesson 4.4, students use the distributive property to rewrite equations in point-slope form as equivalent equations in intercept form.



Lines of Fit

A major reason for studying equations of lines is to learn to make predictions. If you have several data points that lie on a line and you want to predict where another point (x, y) will lie, you can find the equation of the line and evaluate it to find y for a given x -value, or vice versa.

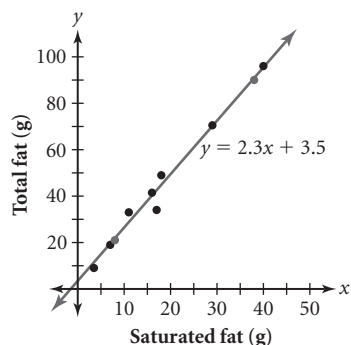
(continued)

Chapter 4 • Fitting a Line to Data (continued)

Most points in a real-world data set don't lie on a single straight line, no matter how linear they look. Measurement error and other "field" factors come into play. So to make predictions, you need to find a line that comes close to the data points. This kind of line is called a *line of fit* for the data. Finding lines of fit gives your student a context to practice finding slopes and equations, and it has useful applications in science and business.

Finding a line of fit by adjusting a calculator graph until the line looks like a good fit gives students experience with relating a line's steepness and y -intercept to its equation. Finding a line of fit by calculating the equation of the line through two particular data points gives students experience with using the slope formula and the point-slope form of a linear equation.

This chapter also shows students how to find a line of fit using Q-points, which builds on the statistics of Chapter 1. Later, in *Discovering Advanced Algebra*, students will learn the method of linear regression to find what's often called the *line of best fit*.



Summary Problem

Here's a table showing the federal minimum wage in various years (taken from Exercise 10 in the Chapter Review). Predict what the minimum wage will be in 2020.

Year	1975	1980	1990
Wage (\$)	2.00	3.10	3.80

Questions you might ask, in your role as student to your student, include:

- If you draw a line that fits the three points well, what would the line's slope and y -intercept be? What prediction would you make?
- If you drew lines through two of the three points, what would the lines' equations be? What predictions would they lead to?
- If you drew a line using Q-points from all the data in Exercise 10, what would the line's equation be? What prediction would it lead to?
- Which, if any, of those lines do you think leads to the best prediction?

Sample Answers

Using different pairs of points will give different lines. For example, the equation for the line between (1975, 2.00) and (1990, 3.80) is approximately $y = 2 + 0.12(x - 1975)$. That equation would predict a value of \$7.40 for 2020. A line of fit for all three points might have approximately the same slope but a slightly higher y -intercept.

To find the equation for the line through the Q-points using all the data from p. 270, you find the quartiles of the x -coordinates and y -coordinates. The first and third quartiles of the x -coordinates are 1976.5 and 1990.5; those of the y -coordinates are about 2.25 and 4.03. So the Q-points that the line of fit will go through are (1976.5, 2.25) and (1990.5, 4.03). The equation for the line through those points has slope $\frac{4.03 - 2.25}{1990.5 - 1976.5}$, or 0.13. The equation $y = 2.25 + 0.13(x - 1976.5)$ gives an estimate of \$7.90 in 2020. Your student may have various reasons to prefer a particular line.

Chapter 4 • Review Exercises

Name _____ Period _____ Date _____

1. (*Lessons 4.1, 4.3*) Consider the line passing through the points (2, 4) and (5, -0.5).
 - a. Find the slope of the line.
 - b. Use the slope to find two other points on the line.
 - c. Write the equation of the line in point-slope form.

2. (*Lesson 4.2*) Sid went for a drive, and after he started he decided to use his trip meter to keep track of the distance he had traveled. He collected the data shown in the table.

Time (h)	Distance (mi)
0.4	75.2
1.5	103.9
1.7	117.3
2.0	130.9
2.7	158.8
3.1	147.0
3.6	179.6
4.9	222.4
5.8	271.3

- a. Plot the data on your calculator. Do the data look approximately linear?
 - b. Look at your scatter plot, and choose two points that seem to be representative of the slope of the data. Find the slope of the line passing through those two points. What is the real-world meaning of the slope in this situation?
 - c. Using the slope you found in 2b, adjust the y -intercept to find a line of fit for the data. What is the real-world meaning of the y -intercept in this situation?
3. (*Lesson 4.4*) Use the distributive property to write an equivalent equation in intercept form.
 - a. $y = 3 + 2(x + 1)$
 - b. $y = 1 + 3(x - 5)$
 - c. $y = -5 - (x - 8)$
4. (*Lesson 4.4*) Factor each expression so that the coefficient of x is 1.
 - a. $4x - 36$
 - b. $-2x + 10$
 - c. $-3x - 15$
 - d. $2x + 7$
5. (*Lessons 4.6, 4.7*) Consider the following data set.

x	1.0	1.7	2.3	3.2	3.5	4.1	4.9	6.2	7.1	7.4
y	35	31	31	28	27	24	25	17	20	18

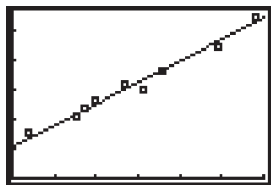
- a. Find the five-number summaries for the x -values and the y -values.
 - b. Create a scatter plot of the data, and determine which Q-points should be used to model this data.
 - c. Find the line of fit based on Q-points for the data. Add the graph of this line to your graph for 5b.

SOLUTIONS TO CHAPTER 4 REVIEW EXERCISES

1. a. -1.5 . To find the slope, divide the difference in y -values by the difference in x -values.

$$\frac{-0.5 - 4}{5 - 2} = \frac{-4.5}{3} = -1.5$$

- b. Answers will vary. Sample answer: Start with the point $(2, 4)$; add 1.5 to the y -value and subtract 1 from the x -value. Repeat this process with the new point you obtain. The resulting points are $(1, 5.5)$ and $(0, 7)$.
- c. Answers will vary. Using the point $(2, 4)$, the equation is $y = 4 - 1.5(x - 2)$. Using the point $(5, -0.5)$, the equation is $y = -0.5 - 1.5(x - 5)$.
2. a. Yes, the data look approximately linear. See the solution for 2c.
- b. Answers will vary. Using the first point and the last point, you get a slope of approximately 36.3. The slope is Sid's average speed in mi/h.
- c. Answers will vary. The line $y = 52.2 + 36.3x$ appears to be a good fit. The y -intercept is an estimate for the reading on the trip meter when Sid started timing his trip.



$[0, 6, 1, 0, 280, 50]$

3. a. $y = 3 + 2(x) + 2(1)$

Use the distributive property.

$$y = 3 + 2x + 2$$

Multiply.

$$y = 5 + 2x$$

Add.

- b. $y = 1 + 3(x) + 3(-5)$

Use the distributive property.

$$y = 1 + 3x + -15$$

Multiply.

$$y = -14 + 3x$$

Add.

- c. $y = -5 - 1(x) - 1(-8)$

Use the distributive property.

$$y = -5 - x + 8$$

Multiply.

$$y = 3 - x$$

Add.

4. a. $4x - 36 = 4(x) - 4(9)$

Factor 4 out of each term.

$$= 4(x - 9)$$

Factor.

- b. $-2x + 10 = -2(x) + -2(-5)$

Factor -2 out of each term.

$$= -2(x - 5)$$

Factor.

- c. $-3x - 15 = -3(x) + -3(5)$

Factor -3 out of each term.

$$= -3(x + 5)$$

Factor.

- d. $2x + 7 = 2(x) + 2(3.5)$

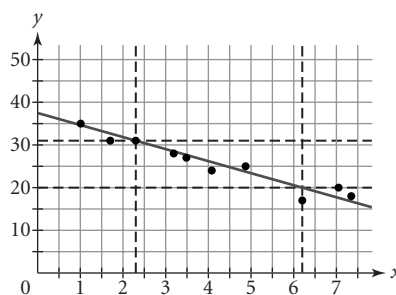
Factor 2 out of each term.

$$= 2(x + 3.5)$$

Factor.

5. a. 1.0, 2.3, 3.8, 6.2, 7.4; 17, 20, 26, 31, 35. See solution to Chapter 1 Exercise 2c in this guide for help in finding five-number summaries.

- b. The Q-points you should use are $(2.3, 31)$ and $(6.2, 20)$. To find the Q-points, graph vertical lines extending from the Q1- and Q3-values for the x -values 2.3 and 6.2, and graph horizontal lines extending from the Q1- and Q3-values for the y -values 20 and 31. This will create a rectangular box. The Q-points are the corners of the box. Use the two Q-points that create the best line of fit for the data, in this case $(2.3, 31)$ and $(6.2, 20)$.



- c. The slope of the line is $\frac{20 - 31}{6.2 - 2.3} = -\frac{11}{3.9} \approx -2.82$. Use the point-slope form of the equation, with the slope and either of the two Q-points found in 5b. The equation is $y = 31 - 2.82(x - 2.3)$ or $y = 20 - 2.82(x - 6.2)$.